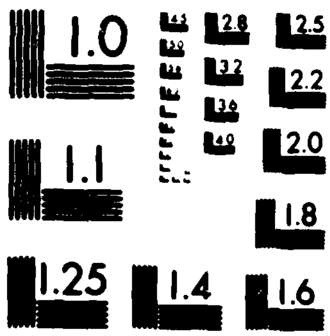
INFERENCE ON PARAMETERS IN A LINEAR MODEL A REVIEW OF 1/1
RECENT RESULTS. (U) PITTSBURGH UNIV PA CENTER FOR
MULTIVARIATE ANALYSIS J MULLER ET AL. JAN 83 TR-83-01
SSIFIED AFOSR-TR-83-0302 F49620-82-K-0001 F/G 12/1 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

INFERENCE ON PARAMETERS IN A LINEAR MODEL
A REVIEW OF RECENT RESULTS

Jochen Müller

C. Radhakrishna Rao

and

Bimal Kumar Sinha

University of Pittsburgh

January 1983

Technical Report No. 83-01

Center for Multivariate Analysis
University of Pittsburgh
Ninth Floor, Schenley Hall
Pittsburgh, PA 15260

This work is sponsored by the Air Force Office of Scientific Research under Contract [redacted]. Reproduction in whole or in part is permitted for any purpose of the United States Government.

F49620-82-K0001

INFERENCE ON PARAMETERS IN A LINEAR MODEL
A REVIEW OF RECENT RESULTS

Jochen Müller

C. Radhakrishna Rao

and

Bimal Kumar Sinha

ABSTRACT

This paper, in three parts, is a review of recent results on inference on parameters in a linear model. In the first part, the Gauss-Markoff theory is extended to the case when the dispersion matrix of the observable random vector is singular. In the second, robustness of inference procedures for departures in the design matrix, the dispersion matrix and distributional assumptions about the error components is considered. Finally, the third part introduces concepts of linear sufficiency and completeness in linear models, without making any distributional assumptions.

Key words and phrases: Gauss-Markoff model, generalized inverse, rank, range of a matrix, minimum variance linear unbiased estimator, least squares theory, generalized projection operator, criterion robustness, inference robustness, likelihood ratio test, uniformly most powerful invariant, spherically symmetric distribution, linear sufficiency, linear completeness.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 83-0302	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) INFERENCE ON PARAMETERS IN A LINEAR MODEL: A REVIEW OF RECENT RESULTS		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL 6. PERFORMING ORG. REPORT NUMBER 83-01
7. AUTHOR(s) Jochen Muller, C. Radhakrishna Rao and Bimal Kumar Sinha		8. CONTRACT OR GRANT NUMBER(s) F49620-82-K-0001
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Multivariate Analysis University of Pittsburgh Pittsburgh PA 15260		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS PE61102F; 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS Mathematical & Information Sciences Directorate Air Force Office of Scientific Research Bolling AFB DC 20332		12. REPORT DATE JAN 83 13. NUMBER OF PAGES 18
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Gauss-Markoff mode; generalized inverse; rank; range of a matrix; minimum variance linear unbiased estimator; least squares theory; generalized projection operator; criterion robustness; inference robustness; likelihood ratio test; uniformly most powerful invariant; spherically symmetric distribution; linear sufficiency; linear completeness.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper, in three parts, is a review of recent results on inference on parameters in a linear model. In the first part, the Gauss-Markoff theory is extended to the case when the dispersion matrix of the observable random vector is singular. In the second, robustness of inference procedures for departures in the design matrix, the dispersion matrix and distributional assumptions about the error components is considered. Finally, the third part introduces concepts of linear sufficiency and completeness in linear models, without making any distributional assumptions.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

83 04 28 048

INFERENCE ON PARAMETERS IN A LINEAR MODEL
A REVIEW OF RECENT RESULTS

Jochen Müller

C. Radhakrishna Rao

and

Bimal Kumar Sinha

Center for Multivariate Analysis
University of Pittsburgh

1. Introduction

We consider the general Gauss-Markoff model

$$Y = X\beta + \epsilon \tag{1.1}$$

where $E(\epsilon) = 0$, $D(\epsilon) = \sigma^2 V$, and the matrices X and V may be singular, and discuss problems of inference on the unknown parameters β and σ^2. We refer to the model (1.1) by the triplet $(Y, X\beta, \sigma^2 V)$. The paper is in three parts. In the first part, the Gauss-Markoff theory is extended to the case when V is singular. In the second, robustness of inference procedures for departures in the design matrix X, the dispersion matrix V and distributional assumptions on Y is considered. The third part introduces the concepts of linear sufficiency and completeness in linear models, without making any distributional assumptions.

The following notations are used throughout the paper.

(i) $\rho(A)$ denotes the rank of a matrix A and $R(A)$, the range of A, i.e., the vector space generated by the columns of A.

(ii) A^- denotes a generalized inverse of A, satisfying the only condition $AA^-A = A$ (see Rao, 1973, p. 34).

(iii) Z denotes a matrix of full rank satisfying the condition $Z'X = 0$, where X is the design matrix.

(iv) The projection operators on $R(A)$ are denoted by (see Rao 1973, p. 48)

$P_{AM} = A(A'MA)^{-}A'M$ where M is p.d. (positive definite)

$P_A = A(A'A)^{-}A'$.

(v) E denotes the expectation operator and D the dispersion operator (providing the variance-covariance matrix of a vector variable).

(vi) For any matrix L, ker L' consists of all vectors a with $L'a = 0$.

(vii) Y: $n\times 1$; X: $n\times m$ with $\rho(X) = r \le m$, β; $m\times 1$.

2. Unified Approach to Linear Estimation

In this section, we consider some methods of estimating the unknown parameters β and σ^2 in the general model (1.1).

2.1 Inverse partitioned matrix approach.

Let

$$\begin{pmatrix} V & X \\ X' & 0 \end{pmatrix}^{-} = \begin{pmatrix} C_1 & C_2 \\ C_3 & -C_4 \end{pmatrix}$$

for any g-inverse. Then the following proposition is proved in Rao (1971).

Proposition 2.1

(i) In the class of linear estimators $L'Y$ such that $X'L = p$, the minimum variance linear unbiased estimator (MVLUE) of $p'\beta$ is $p'\hat{\beta}$ where

$$\hat{\beta} = C_3Y \text{ or } C_2'Y.$$

(ii) If $p'\hat{\beta}$ and $q'\hat{\beta}$ are MVLUE's of $p'\beta$ and $q'\beta$ respectively, then

$$\mathrm{Var}\,(p'\hat{\beta}) = \sigma^2 p'C_4 p$$

$$\mathrm{Cov}\,(p'\hat{\beta}, q'\hat{\beta}) = \sigma^2 p'C_4 q = \sigma^2 q'C_4 p.$$

(iii) An unbiased estimator of σ^2 is

$$\hat{\sigma}^2 = f^{-1}Y'C_4Y, \quad f = \rho(V:X) - \rho(X).$$

2.2 Unified theory of least squares

When V is nonsingular and Y has multivariate normal distribution, we have the following well known results.

(1) Let $\hat{\beta}$ be such that

$$\min_{\beta} (Y-X\beta)'V^{-1}(Y-X\beta) = (Y-X\hat{\beta})'V^{-1}(Y-X\hat{\beta}).$$

Then the MVLUE of $p'\beta$, $p \in R(X')$, is $p'\hat{\beta}$.

(2) $R_0^2 = (Y-X\hat{\beta})'V^{-1}(Y-X\hat{\beta}) \sim \sigma^2\chi^2(f)$

i.e., distributed as χ^2 on f.d.f., where $f = \rho(V:X) - \rho(X)$.

(3) Let $K'\beta = \omega$ be a linear hypothesis where $R(K) \subset R(X')$ and $\rho(K) = h$, and

$$R_1^2 = \min_{K'\beta=\omega} (Y-X\beta)'V^{-1}(Y-X\beta).$$

Then

$$R_1^2 - R_0^2 \sim \sigma^2 \chi^2(h).$$

If V is singular, the above statements are not applicable and the following question arises. Does there exist a symmetric matrix M which takes the place of V^{-1} for which the above properties (1)-(3) hold? The answer is contained in Proposition 2.2 proved in Rao (1973).

<u>Proposition 2.2</u> Let $M = (V + XUX')^-$ for any symmetric g-inverse and U be any symmetric matrix such that $\rho(V;X) = \rho(V+XUX')$.

(i) If $\hat{\beta}$ is such that

$$\min_{\beta} (Y-X\beta)'M(Y-X\beta) = (Y-X\hat{\beta})'M(Y-X\hat{\beta})$$

then the MVLUE of $p'\beta$, $p \in R(X')$, is $p'\hat{\beta}$.

(ii) $R_0^2 = (Y-X\hat{\beta})'M(Y-X\hat{\beta}) \sim \sigma^2\chi^2(f)$, $f = \rho(V:X) - \rho(x)$.

(iii) There is no choice of M for which the property (3) also holds for all testable hypotheses.

Contrary to what is stated in (iii), claims have been made about the existence of M for which the property (3) also holds. This is shown to be not true in Rao (1978).

2.3 Least squares theory with derived restrictions

If V is $n \times n$ and singular, then there exists a matrix N of rank $s = n - \rho(V)$ such that $N'V = 0$ which implies that

$$N'Y - N'X\beta = 0 \text{ w.p.1.} \tag{2.3.1}$$

This stochastic relationship may be considered as a restriction on the parameter β, which is known when Y is observed. In such a case, the following proposition is proved by Goldman and Zelen (1964) and Mitra and Rao (1968).

Proposition 2.3 Let V^- be any g-inverse of V and $\hat{\beta}$ be such that

$$\min_{N'Y=N'X\beta} (Y-X\beta)'V^-(Y-X\beta) = (Y-X\hat{\beta})V^-(Y-X\hat{\beta}) = R_0^2.$$

Then

(i) $p'\hat{\beta}$ is the MVLUE of $p'\beta$, $p \in R(X')$.

(ii) $R_0^2 \sim \sigma^2\chi^2(f)$, $f = \rho(V:X) - \rho(X)$

(iii) If

$$R_1^2 = \min_{\substack{N'Y=N'X\beta \\ K'\beta = w}} (Y-X\beta)V^-(Y-X\beta)$$

then $R_1^2 - R_0^2 \sim \sigma^2\chi^2(h)$, where h is the degrees of freedom of the hypothesis $K'\beta = w$ to be tested. (Note that h is the rank of the variance covariance matrix of $K'\hat{\beta}$ and not necessarily the rank of K.)

2.4 Optimal estimators in a wider class

In Sections 2.1 and 2.2, we considered the class of linear functions of y as estimators ot $p'\beta$, $p \in R(X')$. Now we consider a wider class of functions

$$T(Y) = f(N'Y) + Y'g(N'Y), \tag{2.4.1}$$

where N is as defined in (2.3.1), f is a scalar and g is a vector function, as possible estimators of $p'\beta$. The following proposition is proved in Rao (1979).

<u>Proposition 2.4</u>

(i) $p'\beta$ has an unbiased estimator in the class (2.4.1) iff $p \in R(X')$.

(ii) If $p'\beta$ is unbiasedly estimable, then the MVLUE of $p'\beta$ in the class of (2.4.1) is equivalent w.p. 1 to the MVLUE of $p'\beta$ in the class of linear functions $L'Y$, as considered in Sections 2.1 and 2.2.

(iii) If $L_*'Y$ is the MVLUE of $p'\beta$ in the class $L'Y$, then a general representation of the MVLUE in the wider class (2.4.1) is

$$L_*'Y + f(N'Y) + Y'g(N'Y)$$

where the functions f and g are such that they can be expressed in terms of a function h as

$$f(\xi) = -\xi'h(\xi)$$

$$g(\xi) = N\,h(\xi)$$

for all $\xi \in R(N'X)$ and arbitrary outside $R(N'X)$. Similar approach was given in a paper published later by Harville (1981).

2.5 Generalized projection operator

Consider the general linear model $(Y, X\beta, \sigma^2 V)$, where V may be singular. It is easily seen that

$$Y \in R(V:X) \text{ w.p.1.}$$

The following proposition is established in Rao (1974).

<u>Proposition 2.5</u> Let Z be a matrix of full rank such that $Z'X = 0$. Then:

(i) $R(X)$ and $R(VZ)$ are disjoint, and $R(X:VZ) = R(X:V)$.

(ii) The projection of $y \in R(X:VZ)$ on $R(X)$ along $R(VZ)$ can be expressed as Py where P is any matrix satisfying the conditions

$$PX = X, \quad PVZ = 0.$$

[Such a matrix P is called a generalized projection operator which reduces to the usual projection operator when $\rho(X:VZ) = n$, where n is the order of V. Note that P is not unique when $\rho(X:VZ) < n$].

From the above proposition we deduce:

<u>Proposition 2.6</u> Let P be the projection operator on $R(X)$ along $R(VZ)$ as defined in Proposition 2.5, and CY be an unbiased estimator of $X\beta$ (i.e., $CX = X$). Then

$$D(CY) - D(PY)$$

is non-negative definite, where D denotes the dispersion (variance-covariance) operator, so that PY is the minimum dispersion unbiased estimator of $X\beta$ in the class of linear unbiased estimators.

Note that

$$D(CY) = D(CY-PY+PY)$$
$$= D(CY-PY) + D(PY) + \sigma^2(C-P)VP' + \sigma^2 PV(C-P)'. \qquad (2.5.1)$$

Since $PVZ = 0 \Rightarrow PV = AX'$ for some A, we have

$$(C-P)VP' = (C-P)XA = 0$$

using the conditions $CX = X$ and $PX = X$. Thus from (2.5.1)

$$D(CY) = D(PY) + D(CY-PY)$$

which proves the Proposition 2.6.

Proposition 2.6 answers a question raised by Kempthorne (1976) on the construction of a projection operator when V is singular, and provides a general method for coordinate free estimation through the concept of a projection operator.

From the Proposition 2.6, we have

Proposition 2.7 Let P be the projection operator on R(X) along R(VZ) and $(X'X)^-$ be any g-inverse of X'X. Then

(i) $p'\hat{\beta}$ is the MVLUE of $p'\beta$, $p \in R(X')$, where

$$\hat{\beta} = (X'X)^- X'PY$$

(ii) An unbiased estimator of σ^2 is

$$f^{-1}Y'(I-P')V^-(I-P)Y, \quad f = \rho(V:X) - \rho(X).$$

Reference may also be made to Example 4, Rao (1973, p. 309), where an approach to linear estimation is given without appealing to concepts of linearity, unbiasedness and minimum variance. This is similar to the methods disc ~~ed in ~~ection 4 of this paper.

3. Robustness in the Linear Model

In this Section we will discuss robustness of some statistical procedures in linear models. To be specific, we will be concerned with the robustness of best linear unbiased estimators (BLUEs) in the context of estimation, and likelihood ratio tests in the context of tests of hypotheses when there is specification error in the design matrix and/or in the dispersion matrix. The consequences of deviations from the assumption of normality on tests will also be discussed.

We assume the same set up as in Section 1. Let $(Y, X\beta, \sigma^2 I)$ be the assumed model while $(Y, X\beta, \sigma^2 V)$ be the correct model, resulting in specification error in the dispersion matrix. Throughout this section we assume that V is p.d. Then the BLUE of an estimable linear parametric function $A\beta$ is the same under both the models if and only if

$$A(X'X)^{-}X'VZ = 0 \text{ for all } Z,\ Z'X = 0. \tag{3.1}$$

This follows from the condition that a BLUE must have zero covariance with every error function. Characterization of matrices V satisfying (3.1) is well known [Rao and Mitra (1971); Rao (1967); Zyskind (1967)]. Generally, (3.1) is equivalent to the following representation of V:

$$V = I + X\Lambda_1 X' + Z\Lambda_2 Z' + X\Lambda_4 Z' + Z\Lambda_4' X' \tag{3.2}$$

where Λ_1, Λ_2 and Λ_4 are arbitrary except that $A\Lambda_4 Z' = 0$ and V is p.d. An equivalent representation of V is the following:

$$V = I + X\Lambda_1 X' + Z\Lambda_2 Z' + X_0\Lambda_3 Z' + Z\Lambda_3' X_0' \tag{3.3}$$

where $X_0 = X(I - A^{-}A)$, Λ_1, Λ_2 and Λ_3 are arbitrary except that V is p.d. Some further necessary and sufficient conditions (i.e., equivalent conditions) for the representati n (3.3) to hold are given in the following.

Proposition 3.1. The representation (3.3) is equivalent to any one of the following conditions:

(a) $Z'VZ_1 = 0$

(b) $P_X V^{-1}(I - P_{X_0,V^{-1}})$ is symmetric

(c) $(I - P_{X_0,V^{-1}})(I - P_{X,V^{-1}})$ is symmetric

(d) There exists an orthogonal matrix T such that $T'(I-P_X)T$, $T'(I-P_{X_0})T$, $T'V^{-1}(I-P_{X,V^{-1}})T$ and $T'(I-P_{X_0,V^{-1}})T$ are diagonal matrices.

In the above A: k×m with $\rho(A) = k$, Z_1: n×k is such that $Z_1'Z_1 = I_k$ and $Z_1'Z = 0$ (k×n-p). For a proof of the above proposition, see Mathew and Bhimasankaram (1982). Incidentally, if we demand (3.1) to hold for all A such that $R(A') \subset R(X')$, which means that for every estimable linear parametric function the BLUE is the same under both the models, then we get the following.

Proposition 3.2. (2.1) holds for all A such that $R(A') \subset R(X')$ under any one of the following equivalent conditions:

(a) $X'VZ = 0$

(b) $VX = XQ$ for some Q

(c) VP_X is symmetric

(d) $P_{X,V^{-1}}$ is symmetric.

The result (a) which implies (b) was proved by Rao (1967), and (c) is due to Zyskind (1967). The result (d) appears in Mathew and Bhimansankaram (1982).

Consider next the problem of testing H_0: $A\beta = 0$ assuming normality of the errors. It is well known that the F-test based on $\lambda(X,I) = Y'(I-P_X)I/Y'(I-P_{X_0})Y$ is both LRT and UMPI (under a suitable group of transformations) under the normal model $N(Y, X\beta, \sigma^2 I)$ (see, for example, Lehmann (1959)). We would like to study the robustness properties of this test in so far as whether the properties of its being LRT (criterion robustness) and UMPI (inference robustness) remain valid under deviations from the assumption of normality and the presence of specification errors in the design matrix and/or the disperison matrix.

To begin with, note that if the correct model is $N(Y, X_1\beta, \sigma^2 V)$, denoting by X_1^0 the matrix $X_1(I-A^-A)$, the LRT testing H_0: $A\beta = 0$ is based on

$$\lambda(X_1, V) = Y'V^{-1}(I-P_{X_1^0\, V^{-1}})Y/Y'V^{-1}(I-P_{X_1\, V^{-1}})Y. \tag{3.4}$$

Therefore, the F-test based on $\lambda(X,I)$ under $N(Y, X\beta, \sigma^2 I)$ is LRT under $N(Y, X_1\beta, \sigma^2 V)$ if and only if

$$\lambda(X,I) \equiv \lambda(X_1, V) \text{ for all } Y \tag{3.5}$$

Under the same design matrix $X_1 = X$ but a different dispersion matrix, the condition on the representation of V is the following:

$$V = I + X\Lambda_1 X' + (s-1)ZZ' + X_0\Lambda_3 Z' + Z\Lambda_3' X_0' \tag{3.6}$$

where Λ_1, Λ_3 are arbitrary and s is an arbitrary positive real number subject to (i) V is p.d. and (ii) $Z_1'X\Lambda_1 X'Z_1 = (s-1)I_k$. The following proposition provides other equivalent conditions.

<u>Proposition 3.3</u>. V has the representation (3.6) if and only if any one of the following equivalent conditions holds:

(a) $(I-P_{X_0})V(I-P_{X_0}) = a(I-P_{X_0})$ for some $a > 0$

(b) $V^{-1}(I-P_{X_0 V^{-1}}) = a(I-P_{X_0})$ for some $a > 0$

(c) $\binom{I-P_X}{LP_X}(V-aI)(I-P_X : P_X L') = 0$, for some $a > 0$, with $A = LX$.

Part (c) of this proposition is due to Khatri (1980) and parts (a), (b) are due to Mathew and Bhimasankaram (1982). The representation (3.6) is due to Rao (1967). When V has the intraclass covariance structure, $V = (1-\rho)I + \rho \underset{\sim}{1}\underset{\sim}{1}'$, proceeding directly Ghosh and Sinha (1980) noted that $\lambda(X,I) \equiv \lambda(X,V)$ if and only if $\underset{\sim}{1} \in R(X_0)$. Some generalizations of (3.2) and (3.6) are reported in Chikuse (1981).

Under the same dispersion matrix I but a different design matrix X_1, the F-test remains LRT if and only if $\lambda(X,I) \equiv \lambda(X_1,I)$. This leads to the following.

Proposition 3.4. $\lambda(X,I) \equiv \lambda(X_1,I)$ for all Y if and only if

$$R(X) = R(X_1) \text{ and } R(X_0) = R(X_1^0) \tag{3.7}$$

Finally, the following proposition provides conditions under which (3.5) holds for arbitrary V and X_1.

Proposition 3.5. (3.5) holds if and only if (3.6) and (3.7) hold.

Propositions (3.4) and (3.5) are due to Mathew and Bhimasankaram (1982). The key to all these results, noted earlier by Sinha and Mukhopadhyay (1980), can be stated in the following most general form with a different simpler proof due to Müller and Sinha.

Lemma 3.6: Let A,B,C,D symmetric be such that

$$\frac{y'Ay}{y'By} = \frac{y'Cy}{y'Dy} \text{ for almost all } y.$$

If there is an x such that $Ax = 0$ and $x'Bx \neq 0$ then

$$C = \gamma A \quad \text{for some } \gamma \in \mathbb{R}.$$

Also

$$D = \gamma B$$

provided $A \neq 0$.

<u>Proof</u>: From the assumption it follows immediately that

$$y'Ay\ y'Dy = y'Cy\ y'By \quad \text{for all } y.$$

Especially for $y = x$ this results in $x'Cx = 0$. Now insert $(y + \lambda x)$ to obtain

$$y'Ay[y'Dy + 2\lambda\ y'Dx + \lambda^2\ x'Dx]$$

$$= [y'Cy + 2\lambda\ y'Cx]\ [y'By + 2\lambda\ y'Bx + \lambda^2\ x'Bx].$$

Comparison of the coefficients of λ^3 yields

$$0 = y'Cx\ (x'Bx) \quad \text{for all } y,$$

whence $Cx = 0$. Therefore the coefficients of λ^2 become

$$y'Ay\ x'Dx = y'Cy\ x'Bx$$

from which

$$C = \frac{x'Dx}{x'Bx} A$$

follows. The remainder is evident.

We now turn our attention to the robustness properties of the F-test for non-normal errors. The following result was proved by Ghosh and Sinha (1980).

<u>Proposition 3.7</u>. Let $Y = X\beta + \sigma\underset{\sim}{\epsilon}$ with $\underset{\sim}{\epsilon}$ distributed according to a density $f(\underset{\sim}{\epsilon})$ given by

$$f(\underset{\sim}{\epsilon}) = \int_0^\infty \frac{e^{-\frac{\tau}{2}\underset{\sim}{\epsilon}'\underset{\sim}{\epsilon}}}{(\sqrt{2\pi})^n}\ \tau^{n/2}\ dL(\tau).$$

Then for testing H_0: $A\beta = 0$, the F-test based on $\lambda(X,I)$ is both LRT and UMPI.

Recently Sinha and Drygas (1982) generalized this result to the following.

Proposition 3.8. Let $Y = X\beta + \sigma\epsilon$ with ϵ distributed according to a density $q(\epsilon'\epsilon)$, $q \hat{+}$, convex. Then the F-test based on $\lambda(X,I)$ is both LRT and UMPI.

This result is similar to a robustness property of the Hotelling's T^2-test proved by Kariya (1981) and is based on an application of a representation theorem due to Wijsman (1967). Under a slightly more general distribution of the errors, the following property of a BLUE holds (see Sinha and Drygas (1982)).

Proposition 3.9. Let $Y = X\beta + \sigma\epsilon$ with ϵ having a spherically symmetric distribution. Then for any $c \in \mathbb{R}$ and any n.n.d. matrix C of appropriate order

$$P\{(Gy-A\beta)'C(Gy-A\beta) \leq c^2\} \geq P\{(Ly-A\beta)'C(Ly-A\beta) \leq c^2\}$$

where Gy is any BLUE of estimable $A\beta$ and Ly is any unbiased estimator of $A\beta$.

4. Sufficiency and Completeness in the Linear Model

The well-tried principle of sufficiency has features some of which give rise to a similar concept in the linear model when no distributional assumptions are made. Suppose, for instance, s is a sufficient statistic for some parameter θ and t is independent of s. In this case the expected value of any integrable function h(t) can be written as

$$E_\theta\, h(t) = E(h(t)|s) = \phi(s) \quad \text{a.s.}$$

which is a function independent of θ. Note that $\phi(s)$ must be constant, i.e., t is ancillary, if all underlying distributions share the same null sets (cf. Basu (1958)). It might have been the above equation that led Barnard (1963) to his notion of linear sufficiency. Adjusted to our model $(Y, X\beta, \sigma^2 V)$ it is as follows.

Definition 4.1: A linear statistic Ly is called linearly sufficient if for all linear functions c'y uncorrelated with Ly there is a b such that $E_\beta(c'y) = b'Ly$ a.s.

If V is regular this simply means that the expected value of c'y does not depend on β. Another approach to the idea of linear sufficiency arises from the fact that uniformly minimum variance unbiased estimators are functions of each sufficient statistic. According to this is a definition of Baksalary and Kala (1981), although they originally used a different terminology.

Definition 4.2: A linear statistic Ly is called linearly sufficient if for each linear estimable function p'β the BLUE is a linear function b'Ly of Ly.

On the other hand one may consider that the best prediction of y given any statistic s is the conditional expectation $E_\theta(y|s)$, which is independent of θ if s is sufficient. Reduced to linear terms this property results in a definition that is due to Drygas (1983).

Definition 4.3: A linear statistic Ly is called linearly sufficient if the best linear predictor of y given Ly (written BLP(y|Ly) is independent of β.

If the distribution of y has a density p_θ then, under certain regularity conditions, Fisher's information matrix I is well defined. In this case a statistic s is sufficient if and only if its information matrix I_s equals the original I. In the linear model the assumptions above are met when y is normally distributed and R(X) is contained in R(V). Then the information matrix for the parameter β reads

$$I = \frac{1}{\sigma^2} X' V^- X .$$

This may be regarded as an information measure as well without the normal supposition. One can define therefore:

Definition 4.4: If $R(X) \subset R(V)$ a linear statistic Ly is called linearly sufficient if $I_L = I$.

Each of these definitions can be transformed into algebraic terms, which all turn out to be equivalent. We present two of the handier ones.

Proposition 4.5. Ly is linearly sufficient if and only if $R(X) \subset R(WL')$ or $\ker L \cap R(W) \subset V(\ker X')$. (See Baksalary and Kala (1981), Müller (1982)).

If y is normal with known variance and, in addition, $R(X)$ is a subspace of $R(V)$ then it follows immediately from Definition 3.4 that sufficiency and linear sufficiency are equivalent notions. This attractive property can likewise be confirmed without the regularity condition as was shown by Drygas (1983) and Müller (1982).

Proposition 4.6. If y is normal with known variance then Ly is linearly sufficient if and only if it is sufficient.

But the concept of linear sufficiency also makes some sense without the normal supposition as it becomes evident from Definition 4 .2 and from the following formulation which might be called a linear version of the Rao-Blackwell theorem (see Rao,(1973)): Let Ly be linearly sufficient and $a'\beta$ be any parametric function estimated by $c'y$, say. Then $BLP(c'y|Ly)$ has the same bias as $c'y$ but smaller mean squared error. That means not only BLUEs but all admissible linear estimators are linear functions of Ly. (As for admissibility see Rao (1976).)

Sufficient statistics are especially useful when they are complete. The linear analogue of this concept arises quite naturally.

Definition 4.7: A linear statistic Ly is called linearly complete if each linear function $a'Ly$ that has expected value 0 for all $\beta \in \mathbb{R}^m$ vanishes a.s.

Again the definition can easily be translated into an algebraic expression. Combined with this the above conditions for linear sufficiency turn from inclusions into equations. For normal variables Drygas (1983) showed the accordance with ordinary completeness.

Proposition 4.8. Ly is linearly complete if and only if $R(LV) \subset R(LX)$.

If y is normally distributed this is equivalent to completeness. Ly is linearly sufficient and linearly complete, i.e. sufficient and complete in the normal case with known variance, if and only if $R(X) = R(WL')$ or $\ker L \cap R(W) = V(\ker X')$.

Generally a linearly sufficient Ly does not provide all the information about σ^2 contained in the sample. This deficiency can be compensated when, in addition to Ly, one or more quadratic forms are considered. One would like to extend now the idea of linear sufficiency to this situation, but among the four definitions above only Definition 4.2 can serve this purpose satisfactorily.

Definition 4.9: $(Ly, y'Ty)$ is called quadratically sufficient if Ly is linearly sufficient and the residual sum of squares can be expressed as $y'L'ALy + \alpha\, y'Ty$ for some symmetric A and real α.

Note that the residual sum of squares is a minimum variance unbiased estimator of $\sigma^2 \times$ (degrees of freedom) if y is normal. Things become rather more complicated when one allows for more than one quadratic form while V is singular. With the above definition, however, the following can be proved (cf. Seely (1978), Müller (1982)).

Proposition 4.10. (a) $(Ly, y'Ty)$ is quadratically sufficient if and only if for some $\alpha \in \mathbb{R}$, $\ker L \cap R(W) \subset V(\ker X') \cap \ker X'T \cap \ker (I-\alpha VT)$.

(b) If y is normal a quadratically sufficient $(Ly, y'Ty)$ is sufficient. It is complete if Ly is complete.

(c) If y is normal a sufficient $(Ly, y'Ty)$ is quadratically sufficient provided one of the following two conditions holds.

(i) $y'Ty \geq 0$ a.s. (i.e. WTW is positive semidefinite).

(ii) $y'Ty$ is invariant a.s. (i.e. $WTX = 0$) and Ly is complete (i.e. $R(LV) \subset R(LX)$).

REFERENCES

[1] Baksalary, J.K. and Kala, R. (1981). Linear transformations preserving the best linear unbiased estimator in a general Gauss-Markov model. Ann. Stat. 9, 913-916.

[2] Barnard, G.A. (1963). The logic of least squares. J. Roy. Stat. Soc. B 25, 124-127.

[3] Basu, D. (1958). On statistics independent of a complete sufficient statistic. Sankhyā 15, 377-380 and Sankhyā 20, 223-226.

[4] Chikuse, Y. (1981). Representations of the covariance matrix for robustness in the Gauss-Markov model. Comm. Statist. A10, No. 19, 1997-2004.

[5] Drygas, H. (1983). Sufficiency and completeness in the general Gauss-Markov model. Submitted to Sankhyā.

[6] Ghosh, M. and Sinha, Bimal Kumar (1980). On the robustness of least squares procedures in regression models. J. Mult. Analysis, 10, 332-342.

[7] Goldman, A.J. and Zelen, M. (1964). Weak generalized inverse and minimum variance unbiased estimation. J. Research Nat. Bureau of Standards 68 B, 151-172.

[8] Harville, D.A. (1981). Unbiased and minimum variance unbiased estimation of estimable functions for fixed linear models with arbitrary covariance structure. Ann. Statist. 9, 633-637.

[9] Kariya, T. (1981). A robustness property of Hotelling's T^2-Test. Ann. Statist. 9, 211-214.

[10] Kempthorne, O. (1976). Best linear unbiased estimation with arbitrary covariance matrix. Chapter 14 (pp.203-225) in Essays in Probability and Statistics, Sinko Tsusho Co. Ltd., Tokyo.

[11] Khatri, C.G. (1980). Study of F-tests under dependent model. Sankhyā, Series A, 43, 107-110.

[12] Lehmann, E.L. (1959). Testing Statistical Hypotheses. Wiley, New York.

[13] Mathew, Thomas and Bhimasankaram, P. (1981). On the robustness of the LRT with respect to specification errors in a linear model. Indian Statistical Institute Tech. Report. To appear in Sankhyā.

[14] Mitra, S.K. and Rao, C.R. (1968). Some results in estimation and tests of linear hypotheses under the Gauss-Markoff model. Sankhyā A, 30, 313-322.

[15] Müller, J. (1982). Sufficiency and completeness in the linear model. Tech. Rept. #82-32, Center for Multivariate Analysis, University of Pittsburgh.

[16] Rao, C.R. (1967). Least squares theory using an estimated dispersion matrix and its application to measurement of signals. Proc. Fifth Berkeley Symp. on Math. Stat. and Prob., 1, 355-372.

[17] Rao, C.R. (1971). Unified theory of linear estimation. Sankhyā, A33, 370-396 and Sankhya A36, 447.

[18] Rao, C.R. (1973). Linear Statistical Inference and its Applications, Wiley, New York (First edition, 1965).

[19] Rao, C.R. (1973). Unified theory of least squares. Communications in Statistics, 1, 1-8.

[20] Rao, C.R. (1974). Projectors, generalized inverses and the BLUE's. J. Roy. Statist. Soc.

[21] Rao, C.R. (1976). Estimation of parameters in a linear model. Ann. Statist., 4, 1023-1037.

[22] Rao, C.R. (1978). Least squares theory for possibly singular models. Canad. J. Statist. 3, 105-110.

[23] Rao, C.R. (1979). Estimation of parameters in the singular Gauss-Markoff model. Comm. Statist. A8, 1353-1358.

[24] Rao, C.R. and Mitra, S.K. (1971). Generalized Inverse of Matrices and its Applications, Wiley, New York.

[25] Rao, C.R. and Yanai, H. (1979). General definition and decomposition of projectors and some applications to statistical problems. J. Stat. Planning and Inference 3, 1-17.

[26] Seely, J. (1978). A complete sufficient statistic for the linear model under normality and a singular covariance matrix. Comm. Stat. A7, 1465-1473.

[27] Sinha, Bikas K. and Mukhopadhyay, B.B. (1980). A note on a result of Ghosh and Sinha. Cal. Statist. Assoc. Bull., 29, 169-171.

[28] Sinha, Bimal K. and Drygas, H. (1982). Robustness properties of the F-test and best linear unbiased estimators in linear models. Tech. Rept. 82-28, Center for Multivariate Analysis, University of Pittsburgh.

[29] Wijsman, R.A. (1967). Cross section of orbits and their application to densities of maximal invariants. Fifth Berkeley Symp. Math. Statist. Prob., 1, 389-400.

[30] Zyskind, G. (1967). On canonical forms, nonnegative covariance matrices and best and simple least squares linear estimators in linear models. Ann. Math. Statist. 38, 1092-1109.

END

FILMED

6-83

DTIC